## Course Project Self-Review

Self-Evaluate your course project submission.

1. Does the report include a section describing the data?  
 The summary of the data should help understand the features available and how they will be used for prediction or interpretation.

* No points awarded if there is no summary or it is hard to put together what variables are available or how they might be used. (0 pts)
* One point if there is a basic summary, like a data dictionary. (1 pt)
* One extra point if the summary of the data is presented with graphs of distributions and plots that show the relation between features and the outcome variable. (2 pts)

2. Does the report include a paragraph detailing the main objective of this analysis?

* This report is missing a planning section for the data analysis. (0 pts)
* Yes. This plan includes a detailed subtask section or a good vision of what is possible to do with this data set. (1 pt)
* This plan exceeds expectations. In addition to plan out subtasks and vision for this analysis, it also anticipates possible snags that might be incorporated into preliminary hypothesis of the data. (2 pts)

3. Linear regression models - Does the report include a section with variations of linear regression models and which one is the model that best suits the main objectives of this analysis?

* No. It is not clear if a linear regression was used in this analysis, or the linear regression is missing. (0 pts)
* Yes. At least one regression model is included and it discusses findings and results appropriately. (1 pt)
* Yes, there are at least 2 variations of a regression model. One of them is presented as the better alternative, and some findings are presented. Feedback should include what variations of a regression should be considered (testing splits, cross validation, polynomial features, regularized regressions, etc.). (2 pts)

4. Insights and key findings - Does the report include a clear and well presented section with key findings about the problem and next steps?

* No. There are no takeaways, insights, or findings about this problem. (0 pts)
* Yes. Some takeaways and findings derived from the model are presented (1 pt)
* Yes. Takeaways and findings derived from the model are well presented. The quality of insights or the next steps section award this section an extra point. (2 pts)

5. Next Steps - No model is perfect and it is valuable to highlight aspects of the analysis worth revisiting. Does the report highlight possible flaws in the model and a plan of action to revisit this analysis with additional data or different predictive modeling techniques?

* No. There is no mention of possible flaws or plans to revisit the analysis. (0 pts)
* Yes. There is some discussion presented on possible flaws of this model and a plan to revisit this with additional data or different predictive modeling techniques. (1 pt)
* Yes. There is a comprehensive list of possible flaws of this model and a detailed plan to revisit this with additional data or different predictive modeling techniques. The quality of this section awards it an extra point. (2 pts)

If you scored less than 5 points, please review the lessons taught in this course again.

## End of module review: Regression with Regularization Techniques

### **Regularization Techniques**

Three sources of error for your model are: bias, variance, and, irreducible error.  
  
Regularization is a way to achieve building simple models with relatively low error. It helps you avoid overfitting by penalizing high-valued coefficients. It reduces parameters and shrinks the model.  
  
Regularization adds an adjustable regularization strength parameter directly into the cost function.

Regularization performs feature selection by shrinking the contribution of features, which can prevent overfitting.

In Ridge Regression, the complexity penalty λ is applied proportionally to squared coefficient values.

* The penalty term has the effect of “shrinking” coefficients toward 0.
* This imposes bias on the model, but also reduces variance.
* We can select the best regularization strength lambda via cross-validation.
* It’s a best practice to scale features (i.e. using StandardScaler) so penalties aren’t impacted by variable scale.

In LASSO regression: the complexity penalty λ (lambda) is proportional to the absolute value of coefficients. LASSO stands for : Least Absolute Shrinkage and Selection Operator.

* Similar effect to Ridge in terms of complexity tradeoff: increasing lambda raises bias but lowers variance.
* LASSO is more likely than Ridge to perform feature selection, in that for a fixed λ, LASSO is more likely to result in coefficients being set to zero.

Elastic Net combines penalties from both Ridge and LASSO regression. It requires tuning of an additional parameter that determines emphasis of L1 vs. L2 regularization penalties.  
LASSO’s feature selection property yields an interpretability advantage, but may underperform if the target truly depends on many of the features.  
  
Elastic Net, an alternative hybrid approach, introduces a new parameter α (alpha) that determines a weighted average of L1 and L2 penalties.  
  
Regularization techniques have an analytical, a geometric, and a probabilistic interpretation.